Lung Cancer Detection using ML

Synopsis

Table of Contents

Introduction:

Lung cancer is one of the main cause of death and health issue in many countries with a 5-year survival rate. In this project we use machine learning algorithm to diagnose a cancer and suggest immediate treatment for that in early stages.

In this project we have used KNN and Decision Tree algorithm to predict the accuracy of the cancer. In this project we have used scikit-learn libraries like sklearn and pandas to predict and classify the dataset of lung cancer patients. Slicing the dataset and feature scaling options are used to train the dataset. After that we have used confusion matrix, f1 score and accuracy score to predict the accuracy of the result.

Objective:

Lung Cancer is considered as the deadliest cancer in the worldwide. For this reason many countries are developing strategies for early diagnosis of lung cancer.

In this project our objective is to give best result accuracy of lung cancer patients. To achieve this we use KNeighbors Classifier and Decission Tree algorithm to classify the dataset and give the best accuracy of the result.

Background:

To perform KNN algorithm and Decission Tree algorithm we use scikit-learn library. In background of this project, we used libraries as given below:

Numpy:

NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices.

Pandas:

pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with “relational” or “labeled” data both easy and intuitive.

Sklearn:

It aims to provide simple and efficient solutions to learning problems that are accessible to everybody and reusable in various contexts

KNeighbors Classifier:

It is one of the simplest and widely used classification algorithms in which a new data point is classified based on similarity in the specific group of neighboring data points. This gives a competitive result.

Decission Tree:

A Decision Tree is a Flow Chart, and can help you make decisions based on previous experience.

Hardware and Software Requirements:

Hardware Tools	Minimum Requirements
Processor	I5 or above
Ram	4 GB
HardDisl	50 GB

Software Requirements:

Software Tools	Minimum Requirements
Platform	Windows
Operating System	Windows 10 or avove
Technology	Machine – Learning
Scripting Language	Python
IDE	Google colab

Future Scope:

The lung cancer detection system using machine learning technique is much efficient and gives the best results to the radiologists and assist them. This enhances with additional features for upgrading in future. On this processing system to support radiologist to detect the affected patients as the result.

Conclusion:

We processed the dataset to differentiate the affected patient and level of the growth of cancer by machine learning system. Here it presented an approach to find best accuracy of the cancer result to assist the radiologist and for the future enhancement. Further loads ought to be directed at improving the classifying accuracy levels of result through experiments with various alternatives.

Bibiliography and Reference:

- Dataset: Lung Cancer Dataset(Kaggle.com)
- ML Algorithms(machinelearningmastery.com)
- Scikit-Learn Library(scikit-learn.org)